Genic regions of a large salamander genome contain long introns and novel genes.

Journal: BMC Genomics

Publication Year: 2009

Authors: Jeramiah J Smith, Srikrishna Putta, Wei Zhu, Gerald M Pao, Inder M Verma, Tony Hunter, Susan V Bryant, David M Gardiner, Timothy T Harkins, S Randal Voss

PubMed link: 19144141

Funding Grants: Training in the Biology of Human Enbryonic Stem Cells and Emerging Technologies

Public Summary:
BACKGROUND: The basis of genome size variation remains an outstanding question because DNA sequence data are lacking for organisms with large genomes. Sixteen BAC clones from the Mexican axolotl (Ambystoma mexicanum: c-value = 32 x 10(9) bp) were isolated and sequenced to characterize the structure of genic regions. RESULTS: Annotation of genes within BACs showed that axolotl introns are on average 10x longer than orthologous vertebrate introns and they are predicted to contain more functional elements, including miRNAs and snoRNAs. Loci were discovered within BACs for two novel EST transcripts that are differentially expressed during spinal cord regeneration and skin metamorphosis. Unexpectedly, a third novel gene was also discovered while manually annotating BACs. Analysis of human-axolotl protein-coding sequences suggests there are 2% more lineage specific genes in the axolotl genome than the human genome, but the great majority (86%) of genes between axolotl and human are predicted to be 1:1 orthologs. Considering that axolotl genes are on average 5x larger than human genes, the genic component of the salamander genome is estimated to be incredibly large, approximately 2.8 gigabases! CONCLUSION: This study shows that a large salamander genome has a correspondingly large genic component, primarily because genes have incredibly long introns. These intronic sequences may harbor novel coding and non-coding sequences that regulate biological processes that are unique to salamanders.

Scientific Abstract:
BACKGROUND: The basis of genome size variation remains an outstanding question because DNA sequence data are lacking for organisms with large genomes. Sixteen BAC clones from the Mexican axolotl (Ambystoma mexicanum: c-value = 32 x 10(9) bp) were isolated and sequenced to characterize the structure of genic regions. RESULTS: Annotation of genes within BACs showed that axolotl introns are on average 10x longer than orthologous vertebrate introns and they are predicted to contain more functional elements, including miRNAs and snoRNAs. Loci were discovered within BACs for two novel EST transcripts that are differentially expressed during spinal cord regeneration and skin metamorphosis. Unexpectedly, a third novel gene was also discovered while manually annotating BACs. Analysis of human-axolotl protein-coding sequences suggests there are 2% more lineage specific genes in the axolotl genome than the human genome, but the great majority (86%) of genes between axolotl and human are predicted to be 1:1 orthologs. Considering that axolotl genes are on average 5x larger than human genes, the genic component of the salamander genome is estimated to be incredibly large, approximately 2.8 gigabases! CONCLUSION: This study shows that a large salamander genome has a correspondingly large genic component, primarily because genes have incredibly long introns. These intronic sequences may harbor novel coding and non-coding sequences that regulate biological processes that are unique to salamanders.